# Appendix E
# Monte Carlo Analysis

Variability and uncertainty of the input parameters in a risk assessment represent sources of uncertainty of the assessment. The types of input uncertainty facing the risk assessor include:

1. Natural variability, which is intrinsic in the parameter and cannot be reduced by increased frequency and/or precision in measurement: and

2. Uncertainty of estimation, which can be reduced by increased knowledge about the parameter.

Uncertainty due to natural variability or lack of reliable input data is often so large that the model results are valid only as an indication of the order of magnitude of the output parameter. Sensitivity analysis may be used to identify the input assumptions that have the greatest effects on the model predictions. If those assumptions are highly uncertain, it may be worth investing effort to reduce uncertainty about those assumptions.

Another way to address and quantify uncertainty in modeling is by performing a probabilistic uncertainty analysis. A common method of probabilistic analysis is the Monte Carlo simulation method. In Monte Carlo analysis, the variability and uncertainty of each input parameter is represented by a frequency distribution. The user needs to provide the distribution type (e.g., normal, lognormal, uniform, etc.), along with the mean, standard deviation, and minimum and maximum values of each input parameter, as required by the specific distribution instead of a point estimate as in deterministic analysis. Based on the frequency distribution of the input parameters, the Monte Carlo program selects a randomly generated input data set and calculates the corresponding output. Then, a new input data set is generated at random, and the corresponding new output is calculated. This process is repeated until the statistical distribution of the model output reaches a stable state. In general, for relatively simple physical models and "well behaved" frequency distributions, convergence can be reached in a few thousands of runs. The result of the Monte Carlo analysis is the statistical distribution of the output parameters, with its mean (the arithmetic average), median (the 50th percentile), mode (the most probable value), standard deviation, etc., characterizing the uncertainty of the model predictions. For example, a valid conclusion to draw from the output might be "There is an 8% chance that the exposure dose would exceed the target dose under the conditions evaluated." Probabilistic exposure models can also be used to evaluate food web models as demonstrated in a recent article by MacIntosh, Suter, and Hoffman (1994).

To reduce mistakes and prevent abuses in the use of Monte Carlo techniques in ERAs, good practice principles should be followed. Burmaster and Anderson (1994) proposed 14 principles of good practice to assist people in performing and reviewing probabilistic risk assessments, especially in the context of the Federal and state statutes concerning chemicals in the environments. The authors propose that by following these principles, Monte Carlo risk assessments for hazardous waste sites will be easier to understand, will explicitly distinguish assumptions from data, and will consider effects that could otherwise lead to misinterpretation of the results.

A simplified example of a Monte Carlo simulation using the commercial software, Crystal (BallCB)[1] is presented below. To perform uncertainty analysis with CB, one needs to follow these steps:

1. Develop the model (the risk equation) on an Excel[2] spreadsheet.

2. Identify probability distributions for the input parameters, called Assumptions.

3. Identify the output parameter(s) that need to be analyzed, called Forecast(s).

4. Run the Monte Carlo simulation in Crystal Ball.

5. Stop the simulation when the frequency distribution displayed on-screen for the Forecast is stabilized.

6. Look at the statistics of the Forecast contained in the report generated automatically by Crystal Ball.

7. Modify model and/or input assumptions and rerun until satisfactory results are reached.

The following is an example of an application of Crystal Ball. The calculations simulate a Monte Carlo analysis of deer exposure in a hypothetical contaminated forage browsing scenario.

---

[1] Registered trademark of Decisioneering, Inc.

[2] Registered trademark of Microsoft, Inc.

STEP 1. Develop the mode1 on an Excel spreadsheet.

An Excel spreadsheet was prepared of actual data on forage concentrations over a 12-square-mile core area and a larger 30-square-mile area of potential influence. Forage concentrations were available for each l/4 square mile of the entire site area These chemical concentrations can be displayed as they would appear on a map (Figure E-l), with each cell representing a l/4 square mile of the site.

STEP 2. Identify probability distribution for the input parameters, called Assumptions.

In this case, data on actual forage concentrations are known. An average of all forage concentrations over the 12-square-mile core area or 30-square-mile area of influence would overestimate exposure and risk. Use of single point, hotspot forage concentration would also overestimate risk for a large, mobile foraging deer.

The weekly average forage concentrations a deer may be exposed to provide more valuable information. The minimum weekly foraging area for the deer is 1 square mile. Crystal Ball can generate a distribution of probable l-square-mile forage concentrations by averaging concentrations from randomly selected sets of four contiguous l/4-square-mile areas. It is conservatively assumed that the deer is exposed to at least l/4 square mile of its l-square-mile onsite foraging area within the 12-square-mile core area

STEP 3. Identify the output parameter(s) that need to be analyzed, called Forecast(s):

In the present example, the weekly forage intake was identified as the output parameter.

STEP 4. Run the Monte Carlo simulation in Crystal Ball;

STEP 5. Stop the simulation when the frequency distribution displayed on-screen for the Forecast stabilizes:

After about 1,000 trials the statistical distribution of the model output reached an adequately stable state. The mean, median, mode, standard deviation, variance, skewness, kurtosis, coefficient of variability, range minimum, range maximum, range width, mean standard error, and the frequency chart are summarized in the first part of Table E-1.

Different percentile values for the forecast daily intake are summarized in the second part of Table E-l. In this example, the forage concentrations can be compared with dietary concentrations that are known to be safe.

STEP 6. Look at the statistics for the Forecast contained in the report generated automatically by CB.

STEP 7. Modify mode1 and/or input assumption and rerun until satisfactory results are reached.

Application of the Crystal Ball method to realistic data sets shows that under some circumstances, when hot spot concentrations exceed trigger levels, it can be shown that actual exposures are not likely to exceed trigger levels, and if exceedances occur, they are unlikely to have significant impacts on the exposed population. If risk is excessive, remedial alternatives can be evaluated to see of they achieve acceptable risk levels. For example, a remedial option may be to fence off the “hottest” l/4-square-mile sections. In this model, these hot spots may be blocked off (see Figure E-l) and the exposure levels recalculated.

Note that even quantitative uncertainty analyses have remaining, unquantified uncertainty. In this case, such uncertainties include:

- The validity of assumptions regarding safe forage concentrations.
- The likelihood that deer will forage randomly over the whole site.
- The variability in the measured concentrations due to analysis error, sampling error, or seasonal variations, etc.

Risk assessors should identify these remaining uncertainties to avoid giving the false impression that a quantitative uncertainty analysis fully accounts for all uncertainties. It is unlikely that any model can do that.

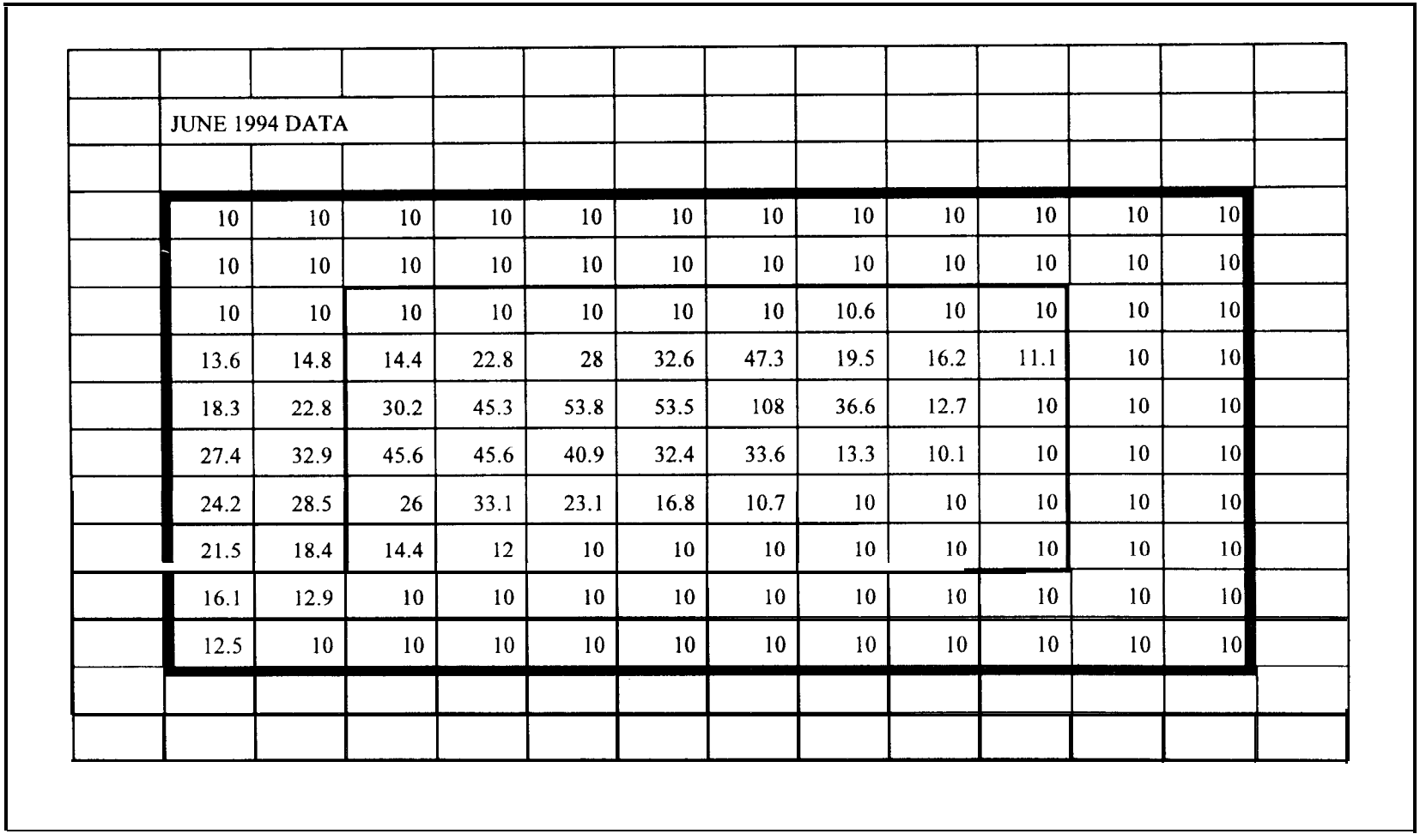


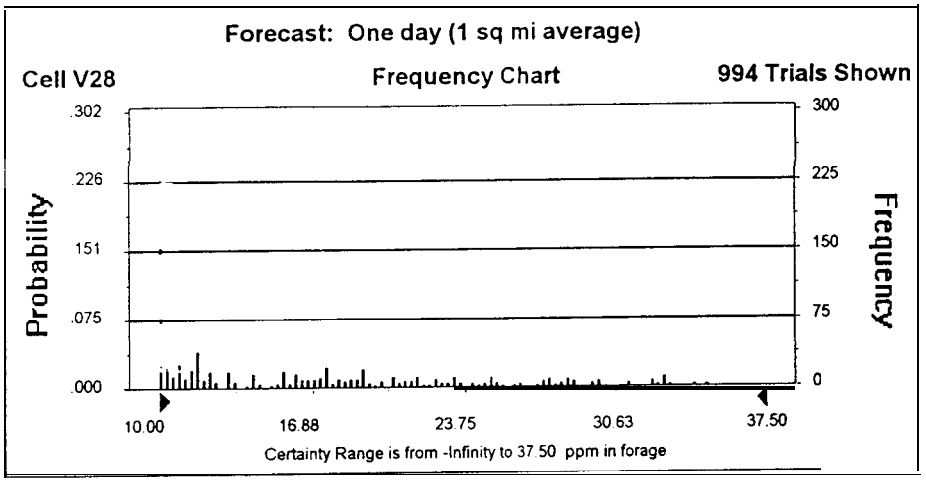

**Figure E-1. Chemical concentrations, ppm in forage, 12-square-mile core area and 30-square-mile area of influence are outlined. Each cell represents the average forage concentration over a 1/4-square-mile forage area.**

Table E-1
Model output

Forecast: One day (1 sq mi average)

Summary:
Certainty Level is 98.95%
Certainty Range is from -Infinity to 37.50 ppm in forage
Display Range is from 10.00 to 37.50 ppm in forage
Entire Range is from 10.00 to 39.93 ppm in forage
After 1,000 Trials, the Std. Error of the Mean is 0.23

| Statistics: | Value |
|---|---|
| Trials | 1000 |
| Mean | 16.59 |
| Median (approx.) | 13.39 |
| Mode (approx.) | 10.15 |
| Standard Deviation | 7.38 |
| Variance | 54.53 |
| Skewness | 0.99 |
| Kurtosis | 2.93 |
| Coeff. of Variability | 0.45 |
| Range Minimum | 10.00 |
| Range Maximum | 39.93 |
| Range Width | 29.93 |
| Mean Std. Error | 0.23 |




Percentiles:

| Percentile | ppm in forage (approx.) |
|---|---|
| 0.0% | 10.00 |
| 2.5% | 10.01 |
| 5.0% | 10.02 |
| 50.0% | 13.39 |
| 95.0% | 32.42 |
| 97.5% | 33.25 |
| 100.0% | 39.93 |

End of Forecast